

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 15, 2002, 04:14:19 ; Search time 3106.14 Seconds
(without alignments)
23342.510 Million cell updates/sec

Title: US-09-652-292-1

Perfect score: 4395

Sequence: 1 gaggggtcttgcagcc.....attatttgtaaaaaaaaaa 4395

Scoring table: OLIGO_NUC

Gapex 60.0 , Gapext 60.0

Searched: 1472140 seqs, 8248589755 residues

Word size : 0

Total number of hits satisfying chosen parameters: 2944280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

GenEmbl:

1: gb_ba:*

2: gb_htg:*

3: gb_in:*

4: gb_om:*

5: gb_ov:*

6: gb_pat:*

7: gb_ph:*

8: gb_pl:*

9: gb_pr:*

10: gb_ro:*

11: gb_sts:*

12: gb_sy:*

13: gb_un:*

14: gb_vl:*

15: em_be:*

16: em_fun:*

17: em_hum:*

18: em_in:*

19: em_om:*

20: em_or:*

21: em_ov:*

22: em_pat:*

23: em_ph:*

24: em_pl:*

25: em_ro:*

26: em_sts:*

27: em_sy:*

28: em_un:*

29: em_vl:*

30: em_htgo_hum:*

31: em_htgo_inv:*

32: em_htgo_rod:*

33: em_htg_hum:*

34: em_htg_inv:*

35: em_htg_rod:*

36: em_htg_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Query	Description
1	4395	100.0	4396	9	AF248053	AF248053 Homo sapi	AF248053 Homo sapi
2	2465	56.1	4167	9	AF321240	AF321240 Homo sapi	AF321240 Homo sapi
3	2424	55.2	4126	9	HS28H201	HS28H201 Novel hum	AL137188 Novel hum
4	1810	41.2	127418	9	HS28H20	HS28H20 Human DNA	AL031055 Human DNA
C 5	182	4.1	385	6	AX150151	AX150151 Sequence	AL015015 Sequence
C 6	156	3.5	411	11	G31518	G31518 human STS S	G31518 human STS S
C 7	99	2.3	1520	9	BC010033	BC010033 Homo sapi	BC010033 Homo sapi
C 8	99	2.3	1524	9	BC005060	BC005060 Homo sapi	BC005060 Homo sapi
C 9	99	2.3	17559	2	AC009133	AC009133 Homo sapi	AC009133 Homo sapi
C 10	99	2.3	181705	2	AC023831	AC023831 Homo sapi	AC023831 Homo sapi
C 11	67	1.5	70876	9	AC023105	AC023105 Homo sapi	AC023105 Homo sapi
C 12	64	1.5	185926	2	AC027632	AC027632 Homo sapi	AC027632 Homo sapi
C 13	63	1.4	108162	2	AC020960	AC020960 Mus muscu	AC020960 Mus muscu
C 14	63	1.4	113982	2	AF192303	AF192303 Homo sapi	AF192303 Homo sapi
C 15	63	1.4	122772	2	AC008874	AC008874 Homo sapi	AC008874 Homo sapi
C 16	63	1.4	164774	2	AC019104	AC019104 Homo sapi	AC019104 Homo sapi
C 17	63	1.4	169241	2	AL445987	AL445987 Homo sapi	AL445987 Homo sapi
C 18	63	1.4	176232	2	AL354659	AL354659 Homo sapi	AL354659 Homo sapi
C 19	63	1.4	198027	2	AC008515	AC008515 Homo sapi	AC008515 Homo sapi
C 20	63	1.4	298734	2	AC008972	AC008972 Homo sapi	AC008972 Homo sapi
C 21	62	1.4	3351	9	AF048726	AF048726 Homo sapi	AF048726 Homo sapi
C 22	62	1.4	75778	9	AL160260	AL160260 Human DNA	AL160260 Human DNA
C 23	62	1.4	150009	2	AC022822	AC022822 Homo sapi	AC022822 Homo sapi
C 24	62	1.4	150856	9	AC087072	AC087072 Homo sapi	AC087072 Homo sapi
C 25	62	1.4	154611	2	AC068893	AC068893 Homo sapi	AC068893 Homo sapi
C 26	62	1.4	166872	2	AC021033	AC021033 Homo sapi	AC021033 Homo sapi
C 27	62	1.4	169725	2	AC073367	AC073367 Homo sapi	AC073367 Homo sapi
C 28	62	1.4	183029	2	AC021030	AC021030 Homo sapi	AC021030 Homo sapi
C 29	62	1.4	186581	2	AC021031	AC021031 Homo sapi	AC021031 Homo sapi
C 30	62	1.4	193877	2	AC009686	AC009686 Homo sapi	AC009686 Homo sapi
C 31	62	1.4	195170	2	AL353576	AL353576 Homo sapi	AL353576 Homo sapi
C 32	62	1.4	196822	2	AC021041	AC021041 Homo sapi	AC021041 Homo sapi
C 33	62	1.4	196933	2	AC079854	AC079854 Homo sapi	AC079854 Homo sapi
C 34	62	1.4	197440	2	AC068580	AC068580 Homo sapi	AC068580 Homo sapi
C 35	62	1.4	208561	2	AC069540	AC069540 Homo sapi	AC069540 Homo sapi
C 36	61	1.4	63545	2	AC087639	AC087639 Homo sapi	AC087639 Homo sapi
C 37	61	1.4	161204	2	AC025772	AC025772 Homo sapi	AC025772 Homo sapi
C 38	61	1.4	166211	2	AC079135	AC079135 Homo sapi	AC079135 Homo sapi
C 39	61	1.4	166263	2	AC019047	AC019047 Homo sapi	AC019047 Homo sapi
C 40	61	1.4	177112	2	AC007498	AC007498 Homo sapi	AC007498 Homo sapi
C 41	61	1.4	185272	2	AC090734	AC090734 Homo sapi	AC090734 Homo sapi
C 42	61	1.4	189036	2	AC019142	AC019142 Homo sapi	AC019142 Homo sapi
C 43	61	1.4	245081	2	AC074138	AC074138 Homo sapi	AC074138 Homo sapi
C 44	60	1.4	85173	9	AL449143	AL449143 Human DNA	AL449143 Human DNA
C 45	60	1.4	98404	2	AL355676	AL355676 Homo sapi	AL355676 Homo sapi

ALIGNMENTS

RESULT 1

AF248053

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

AF248053 4396 bp mRNA 12-APR-2001
Homo sapiens glucose transporter (GLUT10) mRNA, complete cds.

AF248053

AF248053.1 GI:13603726

human.

Homo sapiens

Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 4396)

Fossey,S.C., Mihic,S.J., Craddock,A.L., Mychaleckyj,J.C.,

Dawson,P.A. and Bowden,D.W.

GLUT10: A novel glucose transporter in the type 2 diabetes linked

region of chromosome 20q12-13.1

Unpublished

2 (bases 1 to 4396)

Fossey,S.C., Mihic,S.J., Craddock,A.L., Mychaleckyj,J.C.,

Dawson,P.A. and Bowden,D.W.

TITLE Direct Submission
JOURNAL Submitted (22-MAR-2000) Molecular Genetics, Wake Forest University,
Medical Center Boulevard, Winston-Salem, NC 27106, USA
FEATURES Location/Qualifiers
source 1. 4396
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="20"
/map="20q12-q13.1"
1. 4396
/gene="GLUT10"
251..1876
/gene="GLUT10"
/codon_start=1
/product="glucose transporter"
/protein_id="AAK31911.1"
/db_xref="GI:1360372"
/translation="MGHSPVPLCASVSLGLLTFGYELAVISGALLPLQLDFGLSC
LEQFLVGLSLLGALLASLVGGFLIDCYGRKQAILGSNLVLLAGSTLGLAGSLAWLV
LGRVYGFALISLSSMACCIYVSELVGRQGRVILSYEAGITVGLISYALNVALAGT
PWRHMFWMATAPVQLSLFLPAGTDETHKDLIPLOGGEAPKLPGRPRYSF
LDLFRADNRNRTTVGLGLVLFQOLTGQPNVLCYASTIFSSVGFPHGSSAVLASVL
GAVKAAITLTMGLVDRLRALLAGCALMALSVSGIGLVSPFVMDSPGSLAVPN
ATGOTGLPDGSLDLSPLPIPTNEDQREPLSTAKTKPHRPSGDPSPAPRALIS
SALPGPLPARGHALLRWTLICLAVFSAFSGFQVTLVSEIYPVEIRAFALF
CNSFWAANLFLISLFDLIGTGLSWTLGLTAVLGLGFIYLFVPRTKGOSLAEI
DOQFQRRLTSLFGHRSNLTGIPYSRIEISAAS"
BASE COUNT 939 a 1139 c 1102 g 1216 t
ORIGIN

Query Match 100.0%; Score 4395; DB 9; Length 4396;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 4395; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 gaggggtctctccagggcctgggagcgagggcggtcctgggctccctccgctccg 60
DB 2 GAGGGGCTCTCCAGGCTGGGGCGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 61
QY 61 cctccagggctcgggctgggagcggtggcggtggcggtggcggtggcggtggcg 120
DB 62 CTCTCCAGGCTCGGGGCTGGTGGCGGAGTGGCGCTGGCGGCTGGCGGCGGAGG 121
QY 121 cagggcag 180
DB 122 CAGGGCAG 181
QY 181 ggggactccgggggagatgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 240
DB 182 GGGGACTCCGGGGGAGTCCGGCGCGCGCCCTCAGCGCCCGCCAGCACGCCGCG 241
QY 241 ccgctcgc 300
DB 242 CCGCTCGCCATGGGGACATCCCGACATGTCCTGGCTGGCTGGCTGGCTGGCT 301
QY 301 ggtggcctgaccttggttatgaactggcagtcacatcaggtggcctgacctgacg 360
DB 302 GGTGGCTGACCTTGGTTATGAATGGCAGTCATATCAGGTGGCTGGCTGGCTGG 361
QY 361 cttagcttgggctaaagctggtgagcaggaggttctcctggggcagcgtgctcctg 420
DB 362 CTTGACTTTGGGCTTAAGCTGCTGGAGAGGAGTTCCTGGTGGGCGAGCTGCT 421
QY 421 gctctcctcgcctcctcctcctcctcctcctcctcctcctcctcctcctcctc 480
DB 422 GCTCTCTCGGCTCCTCGTGGTGGCTTCCATGACTGCTATGGCAGGAAGGCC 481
QY 481 atctctgggagcaacttggtgctgctggcagcagcctgacctggcctggcctggt 540
DB 482 ATCTCTGGGAGCAACTTGGTCTGCTGGCAGGACGCTGACCTGGGCTGGCTGG 541
QY 541 ctggcctgagctggtcctgggagcgtggtggtggtggtggtggtggtggtggtg 600

DB 542 CTGGCTGGCTGGTCTCTGGCGCGCTCTGTTGGCTTGGCATTTCCCTCTCCTCCATG 601
QY 601 gcttgctgtatctacgtgtcagagctggtggggccacgagcggggagtgctgctgctc 660
DB 602 GCTTGTGTATCTACGTGTCTCAGAGCTGTGGGGCCACGCGAGCGGGGAGTGTGTGTCTC 661
QY 661 ctctatgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 720
DB 662 CTCTATGAGGAGGAGTACCGTGGGCATCTGCTCTCTCTATGCGCTCAACTATGACTG 721
QY 721 gctggtaacccctgggagtgagggacacatgttcggtgggagcagcagcagcagcagc 780
DB 722 GCTGTACCCCTCTGGGATGGAGGACATGTTGGCTGGGCGACTGCACTTGTCTGTCTG 781
QY 781 caatcctcagcctcctcctcctcctcctcctcctcctcctcctcctcctcctcctc 840
DB 782 CAATCCCTCAGCCT 841
QY 841 ctcatcctcctcagggaggtgagggcccaagctggggccggggagggagggagggaggg 900
DB 842 CTATCCCACTCCAGGAGGTGAGGCCCCAAGCTGGGCGGGGAGGCCACGCTACTCC 901
QY 901 ttctggacctcttccagggcagcgataacatgcagagcgagggcgagggagggaggg 960
DB 902 TTTCTGACCTCTTCAGGGCAGCGGATAACATGCGAGCGGCGGACACAGTGGGCTGGG 961
QY 961 ctggtgctcttcagcaactaacagggcagggcagggcagggcagggcagggcagggc 1020
DB 962 CTGGTGTCTTCCAGCAACTAACAGGGCAGCCCCAACCTGTGTGTGTGTGTGTGTGT 1021
QY 1021 ttcagctccggtgtgttccatggggagcctcagcagcagcagcagcagcagcagcag 1080
DB 1022 TTCAGTCCGCTGGTGTTCATGGGGGATCTCAGCCGCTGTGGCTCTGTGGGGCTTGGC 1081
QY 1081 gcagtgaggtggcagctacccctgacccagccatggggcggtgggtgggtgggtgggt 1140
DB 1082 GCAGTGAAGGTGGCAGCTACCTTGACCGCCATGGGGCTGGTGGAGCGTGGCAGCGCAG 1141
QY 1141 gctctgttactagctggtgtgcccctcagcagcagcagcagcagcagcagcagcagc 1200
DB 1142 GCTCTGTGTCTAGCTGGCTGTGCCCTCATGGCCCTCTCCGCTCAGTGGCATAGGCTCGTC 1201
QY 1201 agcttggcgtgccccctggactcagggcagggcagggcagggcagggcagggcagggc 1260
DB 1202 AGCTTTGGCTGGCCATGGACTCAGGGCCCAAGCTGTCTGGCTGTGCCAATGACCGGG 1261
QY 1261 cagacagggcctcctcctcctcctcctcctcctcctcctcctcctcctcctcctc 1320
DB 1262 CAGACAGGCTCTCCCTGGAGACTTGGCTCTCTGAGGACTCTCTCTCTCTCTCTCTCT 1321
QY 1321 agggacaaatgagacaaagggagcacaatcttgcacgtcctcctcctcctcctcctcct 1380
DB 1322 AGGACCAATGAGGACCAAGAGGAGCCCAATCTTGTCCACCTGCTAAGAAAAACCAAG 1381
QY 1381 ccagatctggagagacccctcagccctcctcctcctcctcctcctcctcctcctcctc 1440
DB 1382 CCCAGATCTGGAGAGCCCTCAGCCCTCTCTGGCTGGCCCTTGAGCTCTGCTCTCTCT 1441
QY 1441 cccctcctcctcctcctcctcctcctcctcctcctcctcctcctcctcctcctcct 1500
DB 1442 CCCCCCTCTGCCCGCTGGGGGATGACATGCTGCGCTGGACCGCATGCTGTGTGTGT 1501
QY 1501 gcttctgctcagtgccctcctcctcctcctcctcctcctcctcctcctcctcctc 1560
DB 1502 GTCTTTGTCTAGTGGCT 1561
QY 1561 atctaccctgtgagatgacaggaagccttcgctccttcgcaacagcttcaactgggg 1620
DB 1562 ATCTACCTGTGGAGATACGAGGAAGAGCTTTCGCTTCTTCTGCAACAGCTTCACT 1621
QY 1621 gccacacctcttcacagcctcctcctcctcctcctcctcctcctcctcctcctcctc 1680
DB 1622 GCCAACCCTCTCATCAGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1681

Qy 1681 accttccctgctctacgaggaactgaccgctgtctctcggcctgggcttccatctctatttattgtt 1740
Db 1682 accttccctgctctacgaggaactgaccgctgtctctcggcctgggcttccatctctatttattgtt 1741
Qy 1741 cctgaacaaagggcctgcttgccagagatagaccagcagcttccaaagagagcgttc 1800
Db 1742 cctgaacaaagggcctgcttgccagagatagaccagcagcttccaaagagagcgttc 1801
Qy 1801 acctgagcttctggccacagcagacccctcctcggcctgcatcccgtagcagccgcagcagatc 1860
Db 1802 acctgagcttctggccacagcagacccctcctcggcctgcatcccgtagcagccgcagcagatc 1861
Qy 1861 tctgggctcctcctgaggaatcctgctgctggaaattctggaaactgtgcttctggcagacc 1920
Db 1862 tctgggctcctcctgaggaatcctgctgctggaaattctggaaactgtgcttctggcagacc 1921
Qy 1921 atctccagcactcctcctcctagcagccagacacagctccagctgcttcttctggagct 1980
Db 1922 atctccagcactcctcctcctagcagccagacacagctccagctgcttcttctggagct 1981
Qy 1981 ggcctcctgcccacaaagctggttcttctgctggggttaaaagagatgaaagcttggagaa 2040
Db 1982 ggcctcctgcccacaaagctggttcttctgctggggttaaaagagatgaaagcttggagaa 2041
Qy 2041 tggccaaattcttctccttcttctcagccctgagggcttcttctgaggaatcctgagcttctcagtc 2100
Db 2042 tggccaaattcttctccttcttctcagccctgagggcttcttctgaggaatcctgagcttctcagtc 2101
Qy 2101 ctgcttcttcccatgacttgacatcttcttctcagcttcttcttataagagaataatctctatgaa 2160
Db 2102 ctgcttcttcccatgacttgacatcttcttctcagcttcttcttataagagaataatctctatgaa 2161
Qy 2161 gctcttcttggcccctggaattcttctcagaagaaatctcaggggtaccaaaccgggagagag 2220
Db 2162 gctcttcttggcccctggaattcttctcagaagaaatctcaggggtaccaaaccgggagagag 2221
Qy 2221 gtttttcccgatatacacccttaaaaccaaagtagagatatcatcttcttctaaatctcttctt 2280
Db 2222 gtttttcccgatatacacccttaaaaccaaagtagagatatcatcttcttctaaatctcttctt 2281
Qy 2281 tcaactgctgaggaacttctcggagggggaagctcttcttcttcttcttcttcttcttcttctt 2340
Db 2282 tcaactgctgaggaacttctcggagggggaagctcttcttcttcttcttcttcttcttcttctt 2341
Qy 2341 ttttttgggtggagctctcattctgttgcagggctggtcctgacttcttggctcactgcacac 2400
Db 2342 ttttttgggtggagctctcattctgttgcagggctggtcctgacttcttggctcactgcacac 2401
Qy 2401 ctccactcctcctgggttcaagagatctcctcctcagcctcctcctcagtagctggattaca 2460
Db 2402 ctccactcctcctgggttcaagagatctcctcctcagcctcctcctcagtagctggattaca 2461
Qy 2461 ggcgctgcccacacacccagcctaatttatttttagcagagatggggttctcagctgttg 2520
Db 2462 ggcgctgcccacacacccagcctaatttatttttagcagagatggggttctcagctgttg 2521
Qy 2521 ggcagctgggtcgtgaaactcctcagctcgaagtgatccacccacccctcagcctccagagctg 2580
Db 2522 gccagctgggtcgtgaaactcctcagctcgaagtgatccacccacccctcagcctccagagctg 2581
Qy 2581 ctgagattacagggccttcttcttcttcttcttcttcttcttcttcttcttcttcttcttcttctt 2640
Db 2582 ctgagattacagggccttcttcttcttcttcttcttcttcttcttcttcttcttcttcttcttctt 2641
Qy 2641 accagaaatttataccctccac 2700
Db 2642 accagaaatttataccctccac 2701
Qy 2701 gctcctgctcgtcgtggtgctcgttaggaatgaccacggcctcagcttcttcttcttcttcttctt 2760
Db 2702 gctcctgctcgtcgtggtgctcgttaggaatgaccacggcctcagcttcttcttcttcttcttctt 2761

Qy 2761 gtataatgggaagcctgtaccaggtcattcttaagaatttctcctgactcagctcagctgagctgg 2820
Db 2762 gtataatgggaagcctgtaccaggtcattcttaagaatttctcctgactcagctcagctgagctgg 2821
Qy 2821 aatctcaaatgctggtcttaggagctgtctccaggaagggtgaggaagtggtcttcttcttcttctt 2880
Db 2822 aatctcaaatgctggtcttaggagctgtctccaggaagggtgaggaagtggtcttcttcttcttctt 2881
Qy 2881 gagatgggttcttggagggcccaaacctgcttctcaatatattgccttcttcttcttcttcttctt 2940
Db 2882 gagatgggttcttggagggcccaaacctgcttctcaatatattgccttcttcttcttcttcttctt 2941
Qy 2941 cttgaaacttgagtaataaacaacacccctgaaacctgaatttctcctcactcagatcaggaatgggga 3000
Db 2942 cttgaaacttgagtaataaacaacacccctgaaacctgaatttctcctcactcagatcaggaatgggga 3001
Qy 3001 taattatgtcccaggggtatatttagaacctgttctccttccagggaggggtcccagctgggt 3060
Db 3002 taattatgtcccaggggtatatttagaacctgttctccttccagggaggggtcccagctgggt 3061
Qy 3061 ccagggcctgggaatttcttcttcttcttcttcttcttcttcttcttcttcttcttcttcttcttctt 3120
Db 3062 ccagggcctgggaatttcttcttcttcttcttcttcttcttcttcttcttcttcttcttcttcttctt 3121
Qy 3121 aagggctcaggggtgaatcagatgggggactgagcaagtagctatgactgagatcagatgataa 3180
Db 3122 aagggctcaggggtgaatcagatgggggactgagcaagtagctatgactgagatcagatgataa 3181
Qy 3181 gaaagggactgacaagaagctcccagatgctgggggagaaaggagagagctaaaaagatccct 3240
Db 3182 gaaagggactgacaagaagctcccagatgctgggggagaaaggagagagctaaaaagatccct 3241
Qy 3241 aggtgctgagctgcttcttcttcttcttcttcttcttcttcttcttcttcttcttcttcttctt 3300
Db 3242 aggtgctgagctgcttcttcttcttcttcttcttcttcttcttcttcttcttcttcttcttctt 3301
Qy 3301 actctgctcctcagcttcttcttcttcttcttcttcttcttcttcttcttcttcttcttcttctt 3360
Db 3302 actctgctcctcagcttcttcttcttcttcttcttcttcttcttcttcttcttcttcttcttctt 3361
Qy 3361 agtgctgagctggtcttcttcttcttcttcttcttcttcttcttcttcttcttcttcttctt 3420
Db 3362 agtgctgagctggtcttcttcttcttcttcttcttcttcttcttcttcttcttcttcttctt 3421
Qy 3421 tacaagcttggttaaacacacagccattataaaaaattaaatgatttataatgatttataatgatt 3480
Db 3422 tacaagcttggttaaacacacagccattataaaaaattaaatgatttataatgatttataatgatt 3481
Qy 3481 aattacattaaaaac 3540
Db 3482 aattacattaaaaac 3541
Qy 3541 ctattctctgcttcttcttcttcttcttcttcttcttcttcttcttcttcttcttcttcttctt 3600
Db 3542 ctattctctgcttcttcttcttcttcttcttcttcttcttcttcttcttcttcttcttcttctt 3601
Qy 3601 acacccgcctccttcttcttcttcttcttcttcttcttcttcttcttcttcttcttcttcttctt 3660
Db 3602 acacccgcctccttcttcttcttcttcttcttcttcttcttcttcttcttcttcttcttcttctt 3661
Qy 3661 aactgctgctggggagctgttcttcttcttcttcttcttcttcttcttcttcttcttcttcttctt 3720
Db 3662 aactgctgctggggagctgttcttcttcttcttcttcttcttcttcttcttcttcttcttcttctt 3721
Qy 3721 tgaattctt 3780
Db 3722 tgaattctt 3781
Qy 3781 caaatataaactttaaagtagcttcttcttcttcttcttcttcttcttcttcttcttcttcttcttctt 3840
Db 3782 caaatataaactttaaagtagcttcttcttcttcttcttcttcttcttcttcttcttcttcttcttctt 3841
Qy 3841 atggagggattattcttctcagtagtgaacactgtctctcctgttctcagctgacagctgctc 3900

Db 2881 TTCTACTTATCTCTCATTTACCCAGGCTCCCTCTTTGGACCTGTAAAGGGTCAGGGTCAAT 2940
 QY 3137 cagatgggagctgagcaaatgactatgactgagatcatgataagaaagggactgacaag 3196
 Db 2941 CAGATGGGGAGCTGAGCAAGTACTATGACTGAGATCATGTAAGGAAGGACTGACAAG 3000
 QY 3197 aagctccagatgctggggagaaatgaagagctaaaaatagatccctaggtgctgagatcctt 3256
 Db 3001 AAGCTCCAGATGCTGGGGAGAAATGAAGAGCTAAATAGATCCTAGTCTGGATGCTTTT 3060
 QY 3257 gtccatccatgctgacacataggggtgctggcagagcccccaaggactctggcctctcgag 3316
 Db 3061 GTATCCATCGTGACATATGGGTGCTGGCAGAGCCCCAAGGACTCTGGCCTCTCGAG 3120
 QY 3317 ttctccatcttctccattctagatgcttcccttgtatccagtgtatccagtgtatgctggagctgac 3376
 Db 3121 TTCTCTATCTTCTCATTTCTAGATGCTTCCCTTGTATCCAGTATGCTGAGGCTGAGTGC 3180
 QY 3377 ttgccaagctgtgagagctggtgctacattttcagagatttttacaagttggttaaca 3436
 Db 3181 TTTGCCAAGCTTGTGAGAGCTGCTTACATTTTCAGGATTTTACAAGTTGGTAACA 3240
 QY 3437 cagccattataaaaaattaaatgatttaattataattataaattacataaaacaa 3496
 Db 3241 CAGCCATTTATAAAAAATTAAATGATTAAATTTAAATTAAGTAAATACATTAACAA 3300
 QY 3497 aaaaattactcaaaattcattacttaatttactacacgttactattctatctgctt 3556
 Db 3301 AAAAAATTACTCAAAATTCATTACTTAATTTACTACCTGTTACTTATTTATCTGCTTT 3360
 QY 3557 tgaggctatttctacataagtaactcttattgagacctaggggagacacgcgcactctctt 3616
 Db 3361 TGAGGCTATTCTTACATAGTAATCTCTTATGGAGAGCTAGGGGAGACACGCGCATCTCTT 3420
 QY 3617 cctgattccccactcaaatgacatcatgcttctgttctgcttgaactgaactgctgggag 3676
 Db 3421 CCTGATTCCTCACTCAATGACATCATGTAGTCTTTGTTGCTTAACTGCTGCTGGGAG 3480
 QY 3677 ttttttgcatacaaaagattagagagactacacatcagggcttgatttattgtttgt 3736
 Db 3481 TGTTTTGTGATCACAAAGATTAGAGAGGACTACATCAGGCTTGATTTATTTGTTGTT 3540
 QY 3737 gatttctagactcagaaatgctgataaaatgctgaatgtaataaattaaactttaaa 3796
 Db 3541 GATTTTCTAGACTTCAGAACATGCTGGATAAATGTCAGTAAATGCAAAATTAACCTTTAA 3600
 QY 3797 gtatgctctgtttgtagccaatatacagtgatagcaccacaaatgagggattattct 3856
 Db 3601 GTATGCTTGTGTGAGCAATACATGCTGTATAGCACCAAAAATGAGGAGTATTCT 3660
 QY 3857 tccagtagtgaacactgtccatccgtttcagctgacagctgctcaaatcatttaagaagg 3916
 Db 3661 TCCAGTAGTTGAACACTGTCTATCCGTTTTCAGCTGACAGCTGCTCAAAATCATTTAAGAAG 3720
 QY 3917 agttcgcactcatttctcattgttttactttgttcttctcactagtgtaaaacaaaat 3976
 Db 3721 AGTTCTGACATTCATTTTCATTTGTTTACTTTTGTCTTCTTCTTCACTAGTGTAAACAAA 3780
 QY 3977 ttcaaccagcattcagtcgcaacacatataccattcttcagtgccctagctgtacagttatc 4036
 Db 3781 TTCAACAGCATTCATGCGCAACCTATAGCCATTTCTCAGTGCCTACTGTCACAGTTATC 3840
 QY 4037 aggaatttttatttcgtagctcaatttttgcataatttcgcaaatcatgcccgaatcgcagtgatgta 4096
 Db 3841 AGGATTTTTTATTTGTAGTCTAAATTTTGTCAAAATCATGGCCAAATCGCAGTATAGTTGA 3900
 QY 4097 ctttggatcaaggttttggcaaaaaaaaataatacaaaaattcttctgtaaaatcaat 4156
 Db 3901 CTTTGGATACAGCTTGGCAAAAAAATAATATACAAAATATTTCTGTAAGAATCAAT 3960
 QY 4157 tggctatagaatttgggataaagaatatttacaataaagaatatttacaataaagaagt 4216
 Db 3961 TGCTATATGGAATTTAGGATAAAGAAATATTTACAATAAAGAAATATTTACAATAAAGAGT 4020

QY 4217 ttattattatttgaagttgtgtgcaacaaacatacccttttattctctgtataaatttatac 4276
 Db 4021 TTATTATTATTGTTGTTGTCGCAACAAACATACCTTTTATCTCTGTAAATTTATAC 4080
 QY 4277 acacaaaaatttaaaaaagattctgtgaagaattattggctatttgaatttagataga 4336
 Db 4081 ACACAAAAATTAAACAAAGATTCTGTAGAAATTAATTGGCTATATGGAATTTAGGATAGA 4140
 QY 4337 atattacaataaaagagattttacaat 4363
 Db 4141 ATATTACAAATAAGAGTATTTACAAT 4167

RESULT 3
 HS28H201
 LOCUS HS28H201 4126 bp mRNA PRI 21-NOV-2000
 DEFINITION Novel human gene mapping to chromosome 20, similar to membrane transporters.
 ACCESSION AL137188
 VERSION AL137188.3 GI:11322734
 KEYWORDS human.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 4126)
 AUTHORS Stavrides,G.S., Hashim,Y., Huckle,E.J. and Deloukas,P.
 TITLE Direct Submission
 JOURNAL Submitted (24-JAN-2000) E-mail contact: humquery@sanger.ac.uk
 COMMENT On Nov 23, 2000 this sequence version replaced gi:11065679.
 This cDNA sequence was assembled from public domain ESTs and single pass sequencing reads from expressed DNA templates, aligned to the genomic DNA sequence from the bacterial clone 28H20 (AL031055). The EST sequences listed match this sequence with an identity of at least 95% between the coordinates shown.
 Further information can be found at
 http://www.sanger.ac.uk/HGP/Chr20/ Sanger Centre name : dJ28H20.C20.1
 Location/Qualifiers
 1. 4126
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="20"
 /map="20q"
 1. 112
 /number=1
 9. 1634
 /codon_start=1
 /product="hypothetical protein"
 /protein_id="CA569822.2"
 /db_xref="GI:11065680"
 /translation="MGHSPVPLPCASVLLGLLTFGYELAVISGALLPLQDLGLSC LQDFIVGSLLLGALLVGGFLIDCYGRKQAILGSLNVLGASLTLGLAGSLAWLV LGRVAVYATISLSMACCIYVSELVGPGRGVLSLYEAGITVGLISLVANVALAGT PWRHMFQWATAPAVIQLSLILFLPAGTDEFATHKDLIPLOGGEAPKLGPRPRYSF LQLFRARDNNRGRTTVGLGLVLFQOLTGCPNVLCYASTIFSSVGFHGGSSAVLSVGL GAVKYAATLTANGLVDNRGRALLAGCALMALSVSGIGLVSFVPMDSGSPCLAVPN ATGQTGLPGDGLLQDSSLPPIPTNEDQREPIILSTAKTKPHRSGDSPAPRLALS SALPGPPLPARGHALLRWLALCMVFSAFSGFGPVTWLVLSLYPVEIRGPAF CSHFNWALNLFISLSFLDLIGTIGLSWTFLLYGLTAVLGLGFIYLPVETKGSIAEI DQFOKRRFTLSFGHRQNSTGIPYSRIEISAAS"
 13. 1296
 /number=2
 1297. 1419
 /number=3
 1338. 1643
 /note="matches EST AA313045"
 1420. 1555
 /number=4
 1556. 4126
 /number=5
 2095. 2137

FEATURES
 source
 exon
 CDS
 exon
 misc_feature
 exon
 exon
 misc_feature

QY	1981	ggccccgccccaaaagtggttttgccttttgtcgggtaaaaggatgaagtgtgaaa	2046
Db	1741	GGCCCCGTGCCCCCAAGGTTGCTCGCTTTTCTGGGTAAAGATGAAGTCTGAGAA	1800
QY	2041	tgcccaattcttcattttgggtttccaagccctgaaggttcttgaggactagttcatgc	2100
Db	1801	TGCCCAACTCTTCATTATTGAGTCTCAGGCCTCGAAGCTTCTCTGAGGATCTAGCTTCATGC	1860
QY	2101	ctcggtttcccattgacctggacatttttgcagttttataagaagaatatctcatgaa	2160
Db	1861	CTCAGTTTCCCATTGACTTGACATCTCTGCAGTATTATATAGAGAATAATTCTCATGAA	1920
QY	2161	gtctttgttgcoccatggaatttttttcaaagaatcctcaggggtaccacaatccggcagggag	2220
Db	1921	GTCTTTGTTCACACATGAGACTTTTTTCTCAAAGAATCTCAAGGTACCAGTCTCGCAGGAA	1980
QY	2221	gttttcccgatatcacccctaataacaaatgagatatcatctttttaaatctctttt	2280
Db	1981	GTCTCTCCCGATATCACCCCTTAATCCAATGAGGATATCATCTTTTCTCAATCTCTTTT	2040
QY	2281	tcaacttgcttgggacatttttcggaaaggggaagtctcttttttactcttatcattttt	2340
Db	2041	TCAACTTGGCTGGGACATTTTCGAAGGGGAAGTCTCTTTTACTCTTATCATTTTTT	2100
QY	2341	ttttttgagtgtagtctcatctctgttgcacagctggcctgatcttggctcactgcaac	2400
Db	2101	TTTTTTGAGTGGAGTCTCATCTGTTGCCAGGCTGGCCTGATCTTGGCTCAGTCAAC	2160
QY	2401	ctcaccttccgttgcacagcaltctctgcctcagcctccttaagtagctggattaca	2460
Db	2161	CTCCACCTCTGAGTTTCAAGCGATTCTTGTGCTCAGCCTCCTAAGCAGTGGGACTACA	2220
QY	2461	ggcgctgcccacacacccagctaatttatatttttagcagagatgggggttcaactgttg	2520
Db	2221	GGCGATGCAACCATACCCAGGCTAATTATTTTATGACAGAGATGGGTTTTCAGTCTGTG	2280
QY	2521	gccagctgtgctgtaactcctgagctcgaagtgtatccaccaacctcagcctcccagatg	2580
Db	2281	GCCAGGCTGGTCGTGAACCTCTGAGTCAAGTGATCCACCCACCTCAGCTCCAGAGTG	2340
QY	2581	ctagattacaagcccttttgactcttttatctgtagltttatgaccccttaattctott	2640
Db	2341	CTAGGATTACAGGCTTTTGACTCTTTATCTGAGTCTTTATGACCCCTCTAAATTCTCTT	2400
QY	2641	accagaatatattatcttccacgacaactcgtactctttgacggagggcctcagttcta	2700
Db	2401	ACCCAGAATATTTATCCTTCACCAACTCTGACTCTTTGACGGAGGCTCTCAGTTCTA	2460
QY	2701	gtccttgctgctgtgtgctatgtctgtaggaatgaccacgggctcctagtttccccattt	2760
Db	2461	GTCTCTGTGCTCTGCTGTCTATGCTGTGTAGGAATGACCACGGGCTCAGTTTCCCATTT	2520
QY	2761	gtataatgggaagcctgtaccaggtctactttaaagatttctcctgactccagtgagctgg	2820
Db	2521	GTATAATGGGAAGCCTGTACCAGGTCAATCTTAGATTTCTCCCTGACTCCAGTGG	2580
QY	2821	aattetaaatgctgtgtcagagctgtctccagatggtgcaggtggcttgcggaag	2880
Db	2581	AATTTCTAAATGCTGTCTPAGGAGTGTCTCCAGGATGCTGCAGGATGGCTTTGCGGAAAG	2640
QY	2881	gagatgggtttggggccaacaaacctgttccaatatgtcctttgcctcttggagacc	2940
Db	2641	CAGATGGGTTTGGAGGCCAACAACTGCTGTCTCAATATTGCCPTTGCCTCTTGGCAGCC	2700
QY	2941	cctgaacttgagtaataacaaactccctgaacctcagtttctctcatctgcagaatgggga	3000
Db	2701	CTTGAACCTGAGTAAAAAACAACCTCCCTGAAACCTCAGTTCTCTCATCTCAGAAATGGGA	2760
QY	3001	taattatgtccaggggtatatttagacctgtttcttccaaaggaggggtccccagctggt	3060
Db	2761	TAAATTATGTCCAGGGGTATATTTAGAACCTGTGTTCTCTTTTCAGAGAGGGTCCCCGTGGT	2820

[illegible]

RESULT	5
AX150151/c	
LOCUS	AX150151 385 bp DNA
DEFINITION	Sequence 126 from Patent WO0136685.
ACCESSION	AX150151
VERSION	AX150151.1 GI:14348179
KEYWORDS	.
SOURCE	human.
ORGANISM	Homo sapiens
	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE	1 (bases 1 to 385)
AUTHORS	Kroes,R.A., Moskai,J.R. and Yamamoto,H.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
 Series: TRAL Plate: 27 Row: c Column: 21
 This clone was selected for full length sequencing because it passed the following selection criteria: Similarity but not identity to protein.

FEATURES

Location/Qualifiers

1. .1520

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="MGC:19694 IMAGE:3506460"

/tissue_type="Kidney, renal cell adenocarcinoma"

/clone_lib="NIH_MGC_14"

/lab_host="DH10B-R"

/note="Vector: pOTB7"

29. .922

/codon_start=1

/product="Unknown (protein for MGC:19694)"

/protein_id="AAH10033.1"

/db_xref="GI:14603130"

/translation="MDAEGIALLPVPTLAALVDSWLRDPCGLNYAALVSGAGPSQA

ALWAKSPCILAGOPFDIAITQLNCQVSWFLPGSKLVPVARVAEVRGPALHLLIGER

VALNTLARGSGIASAAAFAARGAGTGHVACTRTTTCGFRIVEKYLAVGGAASH

RYDLGLVMKDNHVAAGGVEKAVRAQRQAADPALKVECCSSIQEAQVAQAEAGADL

VLLDNFPEELHPATVILKAQPPSVAVASGGITLDNLPPQCGPHIDVISMGMLTQAA

PALDFSLLKFAKEVAPVKTH"

316 a 434 c 491 g 279 t

BASE COUNT

ORIGIN

Query Match 2.3%; Score 99; DB 9; Length 1520;

Best Local Similarity 100.0%; Pred. No. 1e-43;

Matches 99; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2389 gctcactcaacctccactctcgtgggtcaagcgattctcctgcctcagctcctaagta 2448

Db 1444 GCTCACTGCAACCTCCACTTCTGGGTTCAGCGGATTCTCTGCGCTCAGCCCTCTAAGTA 1385.

Qy 2449 gctgggattacaggcgctgcccacacacccagctcaatt 2487

Db 1384 GCTGGGATTACAGGCGGTGCCACACCCAGCTAATT 1346

RESULT

BC005060/c

LOCUS

DEFINITION

BC005060 1524 bp mRNA PRI 12-JUL-2001
Homo sapiens. Similar to quinolinate phosphoribosyltransferase (nicotinate-nucleotide pyrophosphorylase (carboxylating)), clone MGC:12951 IMAGE:2960170, mRNA, complete cds.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

BC005060.1 GI:13477196
human.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

AUTHORS

JOURNAL

Strausberg, R.
Submitted (26-MAR-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA

REMARK

COMMENT

NIH-MGC Project URL: <http://mgc.nci.nih.gov>

Contact: MGC help desk

Email: cgapbs-remail.nih.gov

Tissue Procurement: ATCC

cDNA Library Preparation: Rubin Laboratory

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Genome Sequence Centre,

BC Cancer Agency, Vancouver, BC, Canada

info@bcgsc.bc.ca

Steven Jones, Jennifer Asano, Ian Bosdet, Yaron Butterfield, Susanna Chan, Readman Chiu, Chris Fjell, Erin Garland, Ran Guin, Letticia Hsiao, Martin Krzywinski, Reta Kutsche, Oliver Lee, Soo Sen Lee, Victor Ling, Carrie Mathewson, Candice McLeavy, Steven Ness, Pawan Pandoh, Anna-Lisa Prabhu, Parvaneh Saeedi, Jacqueline Schein, Duane Smailus, Michael Smith, Lorraine Spence, Jeff Stott, Michael Thorne, Miranada Tsai, Nataşja van den Bosch, Jill Vardy, George Yang, Scott Zuyderduyn, Marco Marra.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
 Series: TRAL Plate: 18 Row: a Column: 21.

FEATURES

source

Location/Qualifiers

1. .1524

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="MGC:12951 IMAGE:2960170"

/tissue_type="Muscle, rhabdomyosarcoma"

/clone_lib="NIH_MGC_17"

/lab_host="DH10B-R"

/note="Vector: pOTB7"

28. .921

/codon_start=1

/product="Similar to quinolinate phosphoribosyltransferase (nicotinate-nucleotide pyrophosphorylase (carboxylating))"

/protein_id="AAH05060.1"

/db_xref="GI:13477197"

/translation="MDAEGIALLPVPTLAALVDSWLRDPCGLNYAALVSGAGPSQA

ALWAKSPCILAGOPFDIAITQLNCQVSWFLPGSKLVPVARVAEVRGPALHLLIGER

VALNTLARGSGIASAAAFAARGAGTGHVACTRTTTCGFRIVEKYLAVGGAASH

RYDLGLVMKDNHVAAGGVEKAVRAQRQAADPALKVECCSSIQEAQVAQAEAGADL

VLLDNFPEELHPATVILKAQPPSVAVASGGITLDNLPPQCGPHIDVISMGMLTQAA

PALDFSLLKFAKEVAPVKTH"

321 a 438 c 486 g 279 t

Query Match

Best Local Similarity 2.3%; Score 99; DB 9; Length 1524;

Matches 99; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2389 gctcactcaacctccactctcgtgggtcaagcgattctcctgcctcagctcctaagta 2448

Db 1443 GCTCACTGCAACCTCCACTTCTGGGTTCAGCGGATTCTCTGCGCTCAGCCCTCTAAGTA 1384

Qy 2449 gctgggattacaggcgctgcccacacacccagctcaatt 2487

Db 1383 GCTGGGATTACAGGCGGTGCCACACCCAGCTAATT 1345

RESULT

AC009133/c

LOCUS

DEFINITION

AC009133.6

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

HTG; HTGS_PHASE2; HTGS_DRAFT; HTGS_ACTIVEFIN.

human.

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

AUTHORS

JOURNAL

TITLE

REFERENCE

AUTHORS

JOURNAL

COMMENT

On Apr 25, 2001 this sequence version replaced gi:7689951.

-----Genome Center

Center: Joint Genome Institute
 Center Code: JGI
 Web site: <http://www.jgi.doe.gov>

 Project Information
 Center Project Name: 600095
 Center clone name: RPCI-11_50412

Summary Statistics
 Consensus quality: 172472 bases at least Q40
 Consensus quality: 174610 bases at least Q30
 Consensus quality: 175032 bases at least Q20
 Estimated insert size: 190110; agarose-fp estimation
 Quality coverage: 10.66 in Q20 bases; agarose-fp estimation
 Quality coverage: 11.57 in Q20 bases; sum-of-contigs estimation.
 * NOTE: This is a 'working draft' sequence. It currently
 * consists of 5 contigs. Gaps between the contigs
 * are represented as runs of N. The order of the pieces
 * is believed to be correct as given, however the sizes
 * of the gaps between them are based on estimates that have
 * provided by the submitter.
 * This sequence will be replaced
 * by the finished sequence as soon as it is available and
 * the accession number will be preserved.

1
 22409: contig of 22409 bp in length
 22509: gap of unknown length
 22510: 31954: contig of 9445 bp in length
 31955: 32054: gap of unknown length
 32055: 116297: contig of 84243 bp in length
 116298: 116397: gap of unknown length
 116398: 145562: contig of 29165 bp in length
 145563: 145562: gap of unknown length
 145563: 175599: contig of 29937 bp in length.

FEATURES

Source

1. 175599
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="16"
 /clone="RP11-50412"
 /clone_lib="RPCI human BAC library 11"
 43184 a 44157 c 44428 g 43430 t 400 others

BASE COUNT

ORIGIN

Query Match 2.3%; Score 99; DB 2; Length 175599;
 Best Local Similarity 100.0%; Pred. No. 1.7e-43;
 Matches 99; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2389 gctcactgcaacctcctcctggttcaagcgattctcgtcgtcagctcctctaagta 2448
 |||||
 Db 6908 GCTCAGTGAACCTCCACTTCTGGGTTCAAGCGATTCTCTGCGCTCCTTAAGTA 6849

QY 2449 gctgggattacaggcgctggccaccacacccagctaatt 2487
 |||||

Db 6848 GCTGGGATTACAGCGCGGTGCCACACACCCAGCTAATT 6810
 |||||

RESULT 10

AC023831

LOCUS

AC023831 181705 bp DNA HTG 25-APR-2001
 HOMO sapiens chromosome 16 clone RP11-74E23, WORKING DRAFT
 SEQUENCE, 8 unordered pieces.

ACCESSION

AC023831

VERSION

AC023831.7

KEYWORDS

HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_ACTIVEFIN.

SOURCE

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 181705)

DOE Joint Genome Institute.

Sequencing of Human Chromosome 16

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Unpublished

2 (bases 1 to 181705)

DOE Joint Genome Institute.

Direct Submission

Submitted (18-FEB-2000) Production Sequencing Facility, DOE Joint
 Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
 On Apr 25, 2001 this sequence version replaced gi:11178070.

-----Genome Center

Center: Joint Genome Institute

Center Code: JGI

Web site: <http://www.jgi.doe.gov>

Project Information

Center Project Name: 0

Center clone name: RPCI-11_74E23

Summary Statistics

Consensus quality: 169715 bases at least Q40

Consensus quality: 174817 bases at least Q30

Consensus quality: 176639 bases at least Q20

Estimated insert size: 217000; agarose-fp estimation

Estimated insert size: 181005; sum-of-contigs estimation

Quality coverage: 9.4 in Q20 bases; agarose-fp estimation

Quality coverage: 11.27 in Q20 bases; sum-of-contigs estimation

* NOTE: This is a 'working draft' sequence. It currently
 * consists of 8 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.

1 1212: contig of 1212 bp in length

1213 1312: gap of unknown length

1313 2398: contig of 1086 bp in length

2399 2498: gap of unknown length

2499 3678: contig of 1180 bp in length

3679 3778: gap of unknown length

3779 7829: contig of 4051 bp in length

7830 7929: gap of unknown length

7930 27212: contig of 19283 bp in length

27213 27312: gap of unknown length

27313 49715: contig of 22403 bp in length

49716 49815: gap of unknown length

49816 88863: contig of 39048 bp in length

88864 88963: gap of unknown length

88964 181705: contig of 92742 bp in length.

Location/Qualifiers

1. 181705

/organism="Homo sapiens"

/db_xref="taxon:9606"

/chromosome="16"

/clone="RP11-74E23"

/clone_lib="RPCI human BAC library 11"

BASE COUNT 45562 a 43824 c 45067 g 45113 t 2139 others
 ORIGIN

Query Match

2.3%; Score 99; DB 2; Length 181705;

Best Local Similarity 100.0%; Pred. No. 1.7e-43;

Matches 99; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2389 gctcactgcaacctcctcctggttcaagcgattctcgtcgtcagctcctctaagta 2448
 |||||Db 42813 GCTCAGTGAACCTCCACTTCTGGGTTCAAGCGATTCTCTGCGCTCCTTAAGTA 42872
 |||||QY 2449 gctgggattacaggcgctggccaccacacccagctaatt 2487
 |||||Db 42873 GCTGGGATTACAGCGCGGTGCCACACACCCAGCTAATT 42911
 |||||

RESULT 11

AC023105

LOCUS AC023105 70876 bp DNA PRI 09-MAY-2001
 DEFINITION Homo sapiens BAC clone GSI-13M19 from 7, complete sequence.
 ACCESSION AC023105
 VERSION AC023105.7 GI:11415202
 KEYWORDS HTG.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 70876)
 AUTHORS Sulston, J.E. and Waterston, R.
 TITLE Toward a complete human genome sequence
 JOURNAL Genome Res. 8 (11), 1097-1108 (1998)
 MEDLINE 99063792
 REFERENCE 2 (bases 1 to 70876)
 AUTHORS Mulvane, E., Nguyen, C., Strommatt, C. and Waligorski, J.
 TITLE The sequence of Homo sapiens BAC clone GSI-13M19
 JOURNAL Unpublished
 REFERENCE 3 (bases 1 to 70876)
 AUTHORS Waterston, R.H.
 TITLE Direct Submission
 JOURNAL Submitted (08-FEB-2000) Genome Sequencing Center, Washington University School of Medicine, 4444 Forest Park Parkway, St. Louis, MO 63108, USA
 REFERENCE 4 (bases 1 to 70876)
 AUTHORS Waterston, R.H.
 TITLE Direct Submission
 JOURNAL Submitted (29-NOV-2000) Genome Sequencing Center, Washington University School of Medicine, 4444 Forest Park Parkway, St. Louis, MO 63108, USA
 REFERENCE 5 (bases 1 to 70876)
 AUTHORS Waterston, R.
 TITLE Direct Submission
 JOURNAL Submitted (09-MAY-2001) Department of Genetics, Washington University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
 COMMENT On Nov 29, 2000 this sequence version replaced gi:9838027.
 ----- Genome Center
 Center: Washington University Genome Sequencing Center
 Center code: WUGSC
 Web site: <http://genome.wustl.edu/gsc>
 Contact: sapiens@watson.wustl.edu
 ----- Summary Statistics
 ----- Center project name: H_GS013M19

 NOTICE: This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions.
 This sequence was finished as follows unless otherwise noted: all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.
 MAPPING INFORMATION:
 The sequence of this clone was established as part of a mapping and sequencing collaboration between the NIGRI Chromosome 7 Mapping Project (Eric D. Green, Director), John D. McPherson in the Department of Genetics (Washington University), and the Washington University Genome Sequencing Center. For additional information about the map position of this sequence, see <http://www.nhgri.nih.gov/DIR/OTB/CHR7>, send <mailto:edgreen@nigri.nih.gov>, or see <http://genome.wustl.edu/gsc>
 SOURCE INFORMATION:
 This clone is from the first BAC library from Genome Systems, Inc. (<http://www.genomesystems.com>).
 Cell line: lymphoblastoid
 Haplotypes: two

VECTOR: pBelOBAC
 SELECTION: chloramphenicol
 NEIGHBORING SEQUENCE INFORMATION:
 The clone sequenced to the right is GSI-96J14, 200 bp overlap.
 Actual start of this clone is at base position 1 of GSI-13M19;
 actual end is at base position 63216 of GSI-96J14.
 The clone GSI-13M19 contains a transposon which is not part of the submitted sequence.
 FEATURES
 source 1..70876
 Location/Qualifiers
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="7"
 /map="7"
 /clone="GSI-13M19"
 /clone_lib="GSBAC1"
 134..433
 /rpt_family="Alu"
 repeat_region
 529..810
 /rpt_family="Alu"
 repeat_region
 1312..1585
 /rpt_family="Alu"
 repeat_region
 1589..1884
 /rpt_family="Alu"
 repeat_region
 1995..2125
 /rpt_family="Alu"
 repeat_region
 2322..2407
 /note="similar to EST BE796113 (NID:g10217311)"
 misc_feature
 2325..2434
 /note="similar to EST AI663572 (NID:g4767155) u198c02.y1"
 misc_feature
 2334..2407
 /note="similar to EST AI196214 (NID:g3748820) u170d05.y1"
 misc_feature
 2337..2404
 /note="similar to EST AAS89188 (NID:g2405089) vk08f11.rl"
 misc_feature
 2338..2404
 /note="similar to EST AW763036 (NID:g7694906) ur68c05.y1"
 misc_feature
 2338..2404
 /note="similar to EST BF233698 (NID:g11144387)"
 misc_feature
 2342..2404
 /note="similar to EST AW761967 (NID:g7693882) ur52a03.y1"
 misc_feature
 2342..2404
 /note="similar to EST BF233846 (NID:g11144691)"
 misc_feature
 2342..2404
 /note="similar to EST BF234140 (NID:g11145301)"
 misc_feature
 2344..2404
 /note="similar to EST BF164569 (NID:g11044894)"
 misc_feature
 2345..3115
 /note="similar to EST BF165458 (NID:g11045810)"
 misc_feature
 2380..2525
 /note="similar to EST BE882243 (NID:g10331019)"
 misc_feature
 2382..2525
 /note="similar to EST BE545771 (NID:g9774416)"
 misc_feature
 2382..2525
 /note="similar to EST BE788953 (NID:g10210151)"
 misc_feature
 2382..2525
 /note="similar to EST BE262701 (NID:g9135941)"
 misc_feature
 2384..2525
 /note="similar to EST BE737523 (NID:g10151515)"
 misc_feature
 2385..2525
 /note="similar to EST BE258933 (NID:g9129429)"
 misc_feature
 2385..2525
 /note="similar to EST BF141274 (NID:g10980314)"
 misc_feature
 2386..2525
 /note="similar to EST BE514331 (NID:g9721543)"
 misc_feature
 2389..2525
 /note="similar to EST BE798937 (NID:g10220135)"
 misc_feature
 2391..2525
 /note="similar to EST BF036990 (NID:g10745154)"
 misc_feature
 2392..2525
 /note="similar to EST BE270710 (NID:g9144352)"
 misc_feature
 2392..2525
 /note="similar to EST BF203170 (NID:g11096809)"

misc_feature 2392..2525 /note="similar to EST BF204386 (NID:g11097972)"
misc_feature 2392..2525 /note="similar to EST BF206410 (NID:g11099996)"
misc_feature 2392..2525 /note="similar to EST AV692492 (NID:g10294355)"
misc_feature 2392..2525 /note="similar to EST AW463815 (NID:g7033983)"
misc_feature 2392..2525 /note="similar to EST BE561808 (NID:g9805528)"
misc_feature 2392..2525 /note="similar to EST BE018402 (NID:g8278427) bb79h02.y1"
misc_feature 2392..2525 /note="similar to EST BE737220 (NID:g10151212)"
misc_feature 2392..2525 /note="similar to EST BE903581 (NID:g10394752)"
misc_feature 2392..2525 /note="similar to EST BE513158 (NID:g9720369)"
misc_feature 2392..2525 /note="similar to EST AW247741 (NID:g6590734)"
misc_feature 2392..2525 /note="similar to EST BE561645 (NID:g9805365)"
misc_feature 2392..2525 /note="similar to EST BF205694 (NID:g11099280)"
misc_feature 2392..2525 /note="similar to EST BE785990 (NID:g10207188)"
misc_feature 2392..2525 /note="similar to EST BE616363 (NID:g9897962)"
misc_feature 2392..2525 /note="similar to EST AW248129 (NID:g6591122)"
misc_feature 2392..2525 /note="similar to EST BE798214 (NID:g10219412)"
misc_feature 2392..2525 /note="similar to EST BF203875 (NID:g11097461)"
misc_feature 2392..2525 /note="similar to EST BE798639 (NID:g10219837)"
misc_feature 2392..2525 /note="similar to EST BE208567 (NID:g8751965) ba09f08.y1"
misc_feature 2392..2525 /note="similar to EST BE397389 (NID:g9342754)"
misc_feature 2392..2525 /note="similar to EST AW674129 (NID:g7539364) ba58h03.y1"
misc_feature 2392..2525 /note="similar to EST BE734777 (NID:g10148769)"
misc_feature 2392..2525 /note="similar to EST BE735295 (NID:g10149287)"
misc_feature 2392..2525 /note="similar to EST BE250106 (NID:g9120212)"
misc_feature 2392..2525 /note="similar to EST AW463878 (NID:g7034046)"
misc_feature 2392..2525 /note="similar to EST BE883137 (NID:g10331913)"
misc_feature 2392..2525 /note="similar to EST BE799749 (NID:g10220947)"

Query Match 1.5%; Score 67; DB 9; Length 70876;

Best Local Similarity 100.0%; pred. No. 1.5e-25;

Matches 67; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2376 Tggcctgattgctgactgcaacctccacttctggttcaagcgattctctgcctc 2435
|||||

Db 30679 Tggcctgattgctgactgcaacctccacttctggttcaagcgattctctgcctc 30738
|||||

QY 2436 agcctcc 2442
|||||

Db 30739 AGCCCTCC 30745

RESULT 12

AC027632

LOCUS

DEFINITION

Homo sapiens chromosome 1 clone RP11-487EL, WORKING DRAFT SEQUENCE,

23 unordered pieces.

HTG

DNA

185926 bp

26-AUG-2000

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

JOURNAL

TITLE

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

AC027632
AC027632.6 GI:9929791
HTG: HTGS_PHASE1; HTGS_DRAFT.
human
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
Waterston,R.H.
1 (bases 1 to 185926)
The sequence of Homo sapiens clone
Unpublished
2 (bases 1 to 185926)
Waterston,R.H.
Direct Submission
Submitted (30-MAR-2000) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
On Aug 26, 2000 this sequence version replaced gi:9838225.

----- Genome Center -----

Center: Washington University Genome Sequencing Center

Center code: WUGSC

Web site: http://genome.wustl.edu/gsc/index.shtml

----- Project Information -----

Center project name: H_NH0487E01

----- Summary Statistics -----

Sequencing vector: M13; 95%

Sequencing vector: plasmid; 5%

Chemistry: Dye-terminator Big Dye; 7% of reads

Assembly program: Phrap; version 0.990319

Consensus quality: 172750 bases at least Q40

Consensus quality: 176629 bases at least Q30

Consensus quality: 178424 bases at least Q20

Insert size: 191000; agarose-fp

Insert size: 184940; sum-of-contigs

Quality coverage: 4.41 in Q20 bases; agarose-fp

Quality coverage: 4.66 in Q20 bases; sum-of-contigs

* NOTE: This is a 'working draft' sequence. It currently

* consists of 23 contigs. The true order of the pieces

* is not known and their order in this sequence record is

* arbitrary. Gaps between the contigs are represented as

* runs of N, but the exact sizes of the gaps are unknown.

* This record will be updated with the finished sequence

* as soon as it is available and the accession number will

* be preserved.

* 1

* 1472: contig of 1472 bp in length

* 1473: gap of unknown length

* 1573: contig of 1280 bp in length

* 2853: gap of unknown length

* 2953: contig of 1789 bp in length

* 4742: gap of unknown length

* 4842: contig of 3199 bp in length

* 8041: gap of unknown length

* 8141: contig of 2910 bp in length

* 11051: gap of unknown length

* 11151: contig of 3198 bp in length

* 14349: gap of unknown length

* 14448: contig of 3561 bp in length

* 14499: gap of unknown length

* 18010: contig of 4237 bp in length

* 18110: gap of unknown length

* 22347: gap of unknown length

* 22467: contig of 4817 bp in length

* 22647: gap of unknown length

* 27364: contig of 4051 bp in length

* 27364: gap of unknown length

* 31415: gap of unknown length

* 31514: contig of 4738 bp in length

* 36253: gap of unknown length

* 36353: contig of 5277 bp in length

* 41630: gap of unknown length

* 41729: contig of 4775 bp in length

* 46504: gap of unknown length

* 46605 51108: contig of 4504 bp in length
 * 51109 51208: gap of unknown length
 * 51209 55594: contig of 4386 bp in length
 * 55595 55594: gap of unknown length
 * 55695 61762: contig of 6058 bp in length
 * 61763 61862: gap of unknown length
 * 61863 69971: contig of 8109 bp in length
 * 69972 70071: gap of unknown length
 * 70072 79254: contig of 9183 bp in length
 * 79255 79354: gap of unknown length
 * 79355 94310: contig of 14856 bp in length
 * 94311 94310: gap of unknown length
 * 111262 111362: gap of unknown length
 * 111363 132866: contig of 21504 bp in length
 * 132867 132966: gap of unknown length
 * 132967 159749: contig of 26783 bp in length
 * 159750 159849: gap of unknown length
 * 159850 185926: contig of 26077 bp in length.

FEATURES

source
 1..185926
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="1"
 /clone="RP11-487E1"
 1..1472
 /note="assembly_name:Contig16"
 1573..2852
 /note="assembly_name:Contig17"
 2953..4741
 /note="assembly_name:Contig18"
 4842..8040
 /note="assembly_name:Contig19"
 8141..11050
 /note="assembly_name:Contig20"
 11151..14348
 /note="assembly_name:Contig21"
 14449..18009
 /note="assembly_name:Contig22"
 clone_end:T7
 vector_side:left
 18110..22346
 /note="assembly_name:Contig23"
 22447..27263
 /note="assembly_name:Contig24"
 27364..31414
 /note="assembly_name:Contig25"
 31515..36252
 /note="assembly_name:Contig26"
 36353..41629
 /note="assembly_name:Contig27"
 41730..46504
 /note="assembly_name:Contig28"
 46605..51108
 /note="assembly_name:Contig29"
 51209..55594
 /note="assembly_name:Contig30"
 55695..61762
 /note="assembly_name:Contig31"
 61863..69971
 /note="assembly_name:Contig32"
 70072..79254
 /note="assembly_name:Contig33"
 79355..94210
 /note="assembly_name:Contig34"
 94311..111262
 /note="assembly_name:Contig35"
 111363..132866
 /note="assembly_name:Contig36"
 clone_end:SP6
 vector_side:right
 132967..159749
 /note="assembly_name:Contig37"

misc_feature 159850..185926
 /note="assembly_name:Contig38"
 BASE COUNT 57022 a 36205 c 36048 g 54445 t 2206 others
 ORIGIN
 Query Match 1.5%; Score 64; DB 2: Length 185926;
 Best Local Similarity 100.0%; Pred. No. 8.1e-24;
 Matches 64; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 2380 ctgatttggctcactgcacacctccattctctctgggttcaagcgattctctgctcagcc 2439
 |||||
 DB 82937 CTGATCTGGCTCACTGCAACCTCCACTTCTCTGGTTCAGAGGATTCTCTGCTCAGCC 82996
 |||||
 QY 2440 tctc 2443
 |||||
 DB 82997 TCCT 83000
 |||||
 RESULT 13
 AC020960
 LOCUS
 DEFINITION Mus musculus clone CT7-240L13, WORKING DRAFT SEQUENCE, 1 ordered
 pieces.
 ACCESSION AC020960
 VERSION AC020960.2 GI:9256383
 KEYWORDS HTG: HTGS_PHASE2; HTGS_DRAFT.
 SOURCE house mouse.
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Cranialata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 108162)
 TITLE Sequencing of Mouse
 JOURNAL Unpublished
 2 (bases 1 to 108162)
 DOE Joint Genome Institute.
 Direct Submission
 Submitted (12-JAN-2000) Production Sequencing Facility, DOE Joint
 Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
 On Jul 18, 2000 this sequence version replaced gi:6691270.
 -----Genome Center
 Center: Joint Genome Institute
 Center Code: JGI
 Web site: http://www.jgi.doe.gov

 Project Information
 Center Project Name: 1045618
 Center clone name: RG-MBAC_240L13

 Summary Statistics
 Consensus quality: 103750 bases at least Q40
 Consensus quality: 106923 bases at least Q30
 Consensus quality: 107282 bases at least Q20
 Estimated insert size: 130000; pulse field gel estimation
 Estimated insert size: 108212; sum-of-contigs estimation
 Quality coverage: 6.53 in Q20 bases; pulse field gel estimation
 Quality coverage: 7.85 in Q20 bases; sum-of-contigs estimation.
 * NOTE: This is a 'working draft' sequence. It currently
 * consists of 1 contigs. Gaps between the contigs
 * are represented as runs of N. The order of the pieces
 * is believed to be correct as given, however the sizes
 * of the gaps between them are based on estimates that have
 * provided by the submitter.
 * This sequence will be replaced
 * by the finished sequence as soon as it is available and
 * the accession number will be preserved.
 * 1 108162: contig of 108162 bp in length.
 Location/Qualifiers
 1..108162
 /organism="Mus musculus"
 /db_xref="taxon:10090"
 /clone="CT7-240L13"

BASE COUNT 34815 a 21949 c 21217 g 30178 t 3 others
ORIGIN

Query Match 1.4%; Score 63; DB 2; Length 108162;
Best Local Similarity 100.0%; Pred. No. 2.8e-23;
Matches 63; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2381 tgatttggtcactgcaacctcactctcctggttcaagcattctcctgctcagcct 2440
|||||
Db 101068 TGATCTGGCTCAGTCACTGCACTCTGCTGCTCAAGGATCTCTGCTCAGCCT 101127
QY 2441 cct 2443
|||
Db 101128 CCT 101130

RESULT 14
AF192303
LOCUS
DEFINITION Homo sapiens chromosome 8 clone CTA-204B4 map 8q11-12, WORKING
DRAFT SEQUENCE, 8 unordered pieces.
ACCESSION AF192303.3 GI:14327853
VERSION
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 113982)
AUTHORS Schilhabel,M.B., Baumgart,C., Blechschmidt,K., Dette,M., Jahn,N.,
Lehmann,K., Menzel,U., Polley,A., Reichwald,K., Schudy,A.,
Siddiqui,R., Taudien,S., Wen,G., Korenberg,J., Rosenthal,A. and
Platzer,M.
Chromosome 8 genomic sequence
Unpublished
2 (bases 1 to 113982)
REFERENCE Blechschmidt,K., Schilhabel,M., Schattevoy,R., Dette,M., Menzel,U.,
Korenberg,J.R. and Rosenthal,A.
Direct Submission
Submitted (04-OCT-1999) Genome Analysis, Institute of Molecular
Biotechnology, Beutenbergstrasse 11, Jena 07745, Germany
On Jun 7, 2001 this sequence version replaced gi:8151740.

Center: Institute of Molecular Biotechnology
Center code: IMB
Web site: http://genome.imb-jena.de/
Contact: gscj-submit@genome.imb-jena.de
----- Project Information
Center project name: H211
Center clone name: CTA-204B4
----- Summary Statistics
Sequencing vector: M13; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990329
Consensus quality: 106890 bases at least Q40
Consensus quality: 109881 bases at least Q30
Consensus quality: 111616 bases at least Q20
Quality coverage: 5,74 x in Q20 bases; sum-of-contigs

Sequence Quality Assessment:
This entry has been annotated with sequence quality
estimates computed by the Phrap assembly program.
All manually edited bases have been reduced to quality 10.
Quality levels above 40 are expected to have less than
1 error in 10,000 bp.
Base-by-base quality values are not generally visible from the
GenBank flat file format but are available as part
of this entry's ASN.1 file.

* NOTE: This is a 'working draft' sequence. It currently
* consists of 8 contigs. The true order of the pieces

* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence.
* as soon as it is available and the accession number will
* be preserved.
* 1 3553: contig of 3553 bp in length
* 3554 3653: gap of unknown length
* 3654 8755: contig of 5102 bp in length
* 8756 8855: gap of unknown length
* 8856 35675: contig of 26820 bp in length
* 35676 35775: gap of unknown length
* 35776 85484: contig of 49709 bp in length
* 85485 85585: gap of unknown length
* 85586 94427: contig of 8843 bp in length
* 94428 94527: gap of unknown length
* 94528 105891: contig of 11364 bp in length
* 105892 112771: gap of unknown length
* 112772 112371: contig of 6280 bp in length
* 112372 113982: contig of 1611 bp in length.
* Location/Qualifiers
1. 113982
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="8"
/map="8q11-12"
/clone="CTA-204B4"
BASE COUNT 28505 a 28727 c 27506 g 28544 t 700 others
ORIGIN

FEATURES

source

Query Match 1.4%; Score 63; DB 2; Length 113982;
Best Local Similarity 100.0%; Pred. No. 2.8e-23;
Matches 63; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2412 gggttcaagcattctcctgctcagcctcctaaagtaagtcaggattacagcgcgtgcc 2471
|||||
Db 20488 GGGTCAAGCATTCCTCTGCTCAGCCTCTAAAGTACGTGGATTACAGCGCGTGCCA 20547
QY 2472 cca 2474
|||
Db 20548 CCA 20550

RESULT 15

AC008874

LOCUS

DEFINITION

AC008874 122772 bp DNA HTG 14-FEB-2001

9 ordered pieces.

AC008874

AC008874.5 GI:12830099

HTG; HTGS_PHASE2; HTGS_DRAFT.

human.

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 122772)

DOE Joint Genome Institute.

Sequencing of Human Chromosome 5

Unpublished

2 (bases 1 to 122772)

DOE Joint Genome Institute.

Direct Submission

Submitted (03-AUG-1999) Production Sequencing Facility, DOE Joint

Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA

On Feb 14, 2001 this sequence version replaced gi:9954585.

-----Genome Center

Center: Joint Genome Institute

Center Code: JGI

Web site: http://www.jgi.doe.gov

Project Information

```

Center Project Name: 704688
Center clone name: C17B-H1_220003
-----
Summary Statistics
Consensus quality: 116661 bases at least Q40
Consensus quality: 120243 bases at least Q30
Consensus quality: 121387 bases at least Q20
Estimated insert size: 125000; pulse field gel estimation
Estimated insert size: 121972; sum-of-contigs estimation
Quality coverage: 8.51 in Q20 bases; pulse field gel estimation
Quality coverage: 8.72 in Q20 bases; sum-of-contigs estimation.
* NOTE: This is a 'working draft' sequence. It currently
* consists of 9 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as given, however the sizes
* of the gaps between them are based on estimates that have
* provided by the submitter.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
1
36663 36662: contig of 36662 bp in length
36763 36762: gap of unknown length
45137 45137: contig of 8375 bp in length
45138 45237: gap of unknown length
45238 54176: contig of 8939 bp in length
54177 54276: gap of unknown length
54277 54719: contig of 31443 bp in length
85720 85819: gap of unknown length
85820 88481: contig of 2662 bp in length
88482 88581: gap of unknown length
88582 93330: contig of 4749 bp in length
93331 93430: gap of unknown length
93431 96495: contig of 3065 bp in length
96496 96595: gap of unknown length
96596 119670: contig of 23075 bp in length
119671 119770: gap of unknown length
119771 122772: contig of 3002 bp in length.

FEATURES
    source
    1..122772
        /organism="Homo sapiens"
        /db_xref="taxon:9606"
        /chromosome="5"
        /clone="CTD-220003"
        /clone_lib="Caltech human BAC library D"
BASE COUNT 39486 a 25066 c 23097 g 34312 t 811 others
ORIGIN

Query Match 1.4%; Score 63; DB 2; Length 122772;
Best Local Similarity 100.0%; Pred. No. 2.8e-23;
Matches 63; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2381 tgatcttggtcactgcacactccacttcctcgggttcgaagattctctcgcctcagcct 2440
|||||
Db 120510 TGATCTTGGGCTCACTGCAACCTCCACTTCCTGGGTTCAAGCGATTCTCCTGCTCAGCCT 120569

Qy 2441 cct 2443
|||
Db 120570 CCT 120572

Search completed: March 15, 2002, 07:03:55
Job time: 10176 sec

```

THIS PAGE BLANK (USPTO)